

Strategic Point of Sale Placement

AXEL GOBLET

Problem Description

2

- ▶ Maximizing sales → Maximizing revenue
- ▶ Companies distribute goods through physical points of sales (POSs)
- ▶ Smart placement of POSs could help to increase sales
- ▶ Goal: find surroundings that increase sales of POSs

Available Data

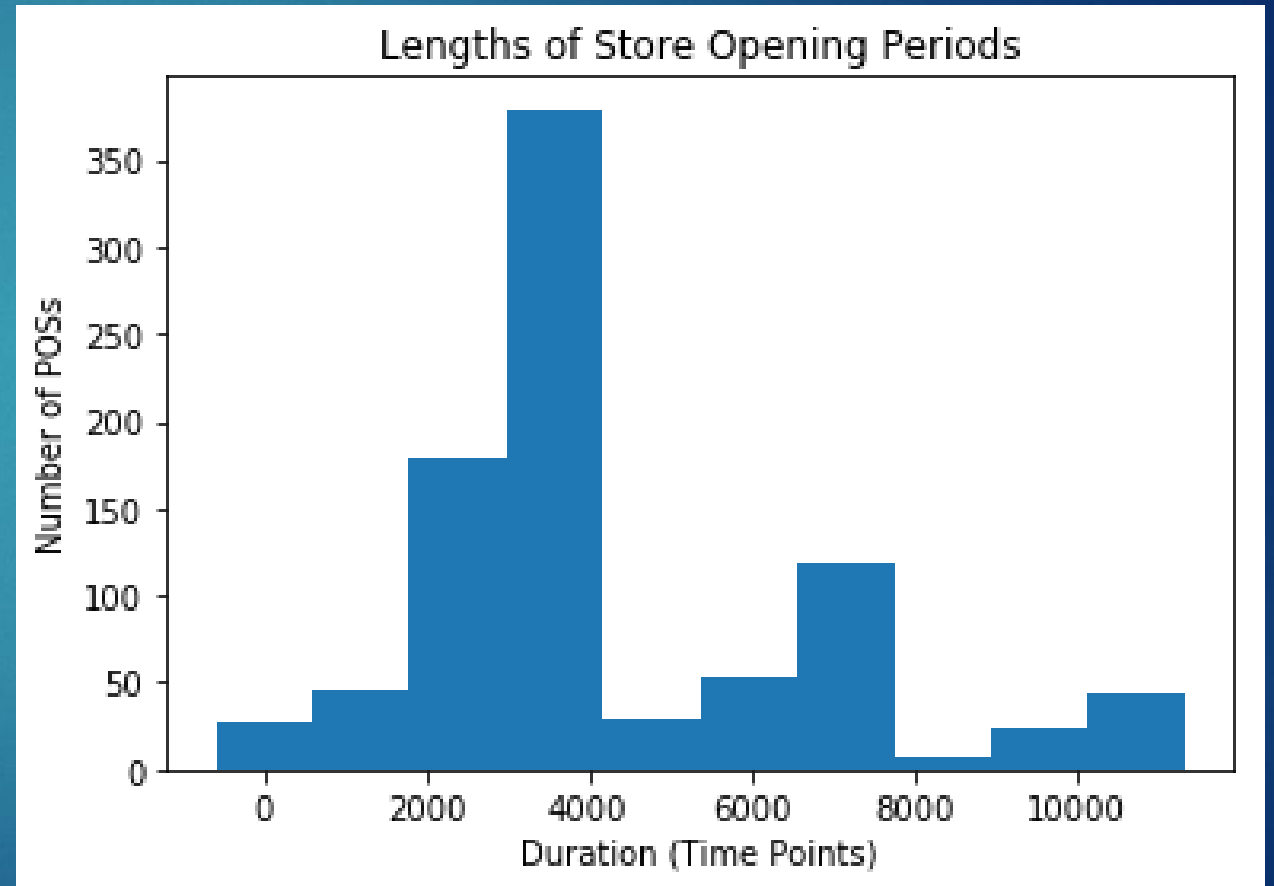
3

- ▶ Sales data for POSs
 - ▶ 906 POSs
 - ▶ Hourly sales data
 - ▶ August 2015 – June 2017
 - ▶ ~12000 time points
- ▶ Information on surroundings of POSs
 - ▶ 546 POSs
 - ▶ 89 types of surroundings (restaurants, embassies, banks,...)
 - ▶ Various characteristics (rating, opening hours,...)

Target Variable

4

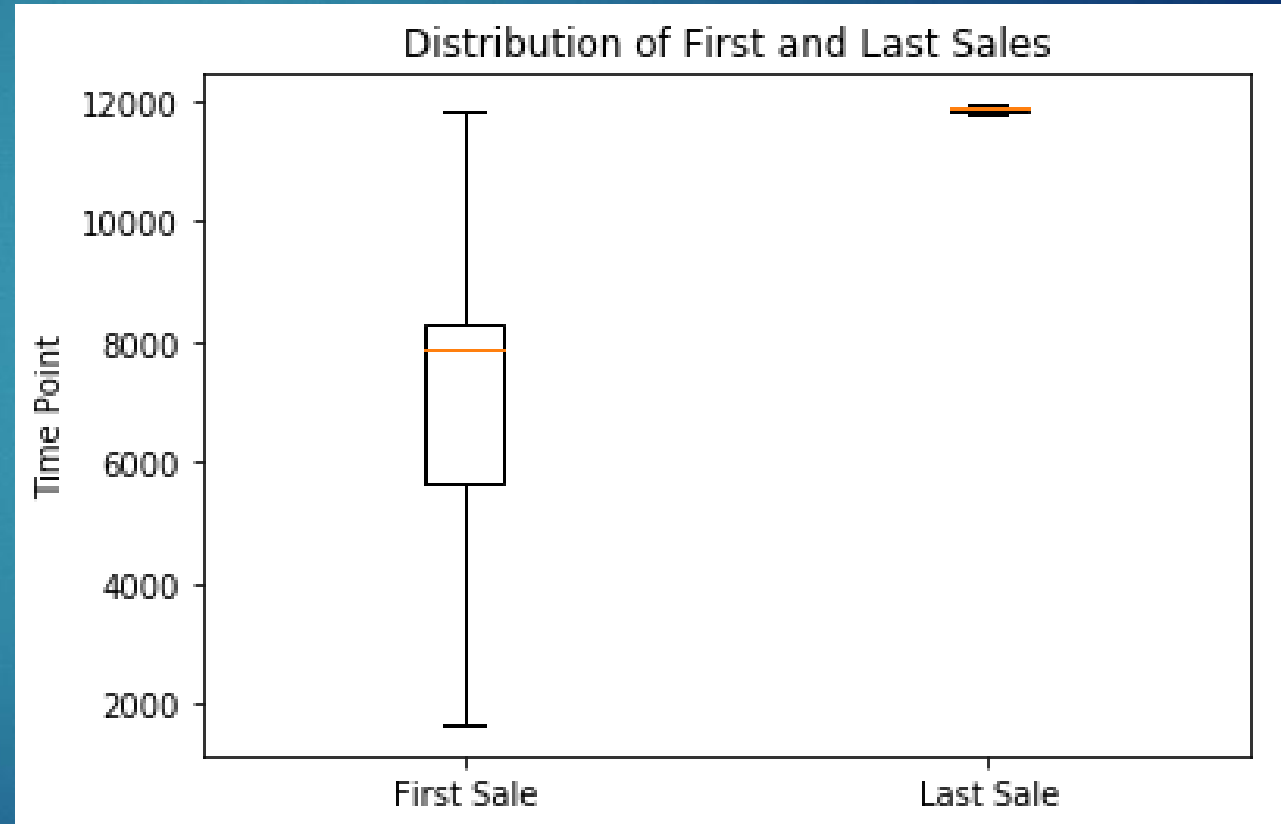
- ▶ POSs are not active over the full period
- ▶ Opening period: first sale until last sale
- ▶ Target variable: average sales during opening period
 - ▶ Regression task



Features (Surroundings)

5

- ▶ First sales of POSs vary
- ▶ Add feature 'first sale' to handle this variety



Features (Surroundings)

6

- ▶ Number of nearby surroundings of each type
- ▶ Number of nearby surroundings
 - ▶ Busy locations
- ▶ Average rating of surroundings of each type
- ▶ Average rating of surroundings
 - ▶ Quality of neighbourhood

Models

7

- ▶ Focus on decision tree-based ensembles
 - ▶ Decision trees → Inherent feature importance & selection
 - ▶ Ensembles → Good predictive power
- ▶ Random Forest
- ▶ AdaBoost

Model Training

8

- ▶ Apply 10-fold cross validation
- ▶ Hand-tweak parameters
 - ▶ Use random search for comparison
- ▶ Performance measure: R^2
 - ▶ Shows relation to baseline (variance of the data)
 - ▶ Punishes outliers → Reduces big errors in predictions

Performance

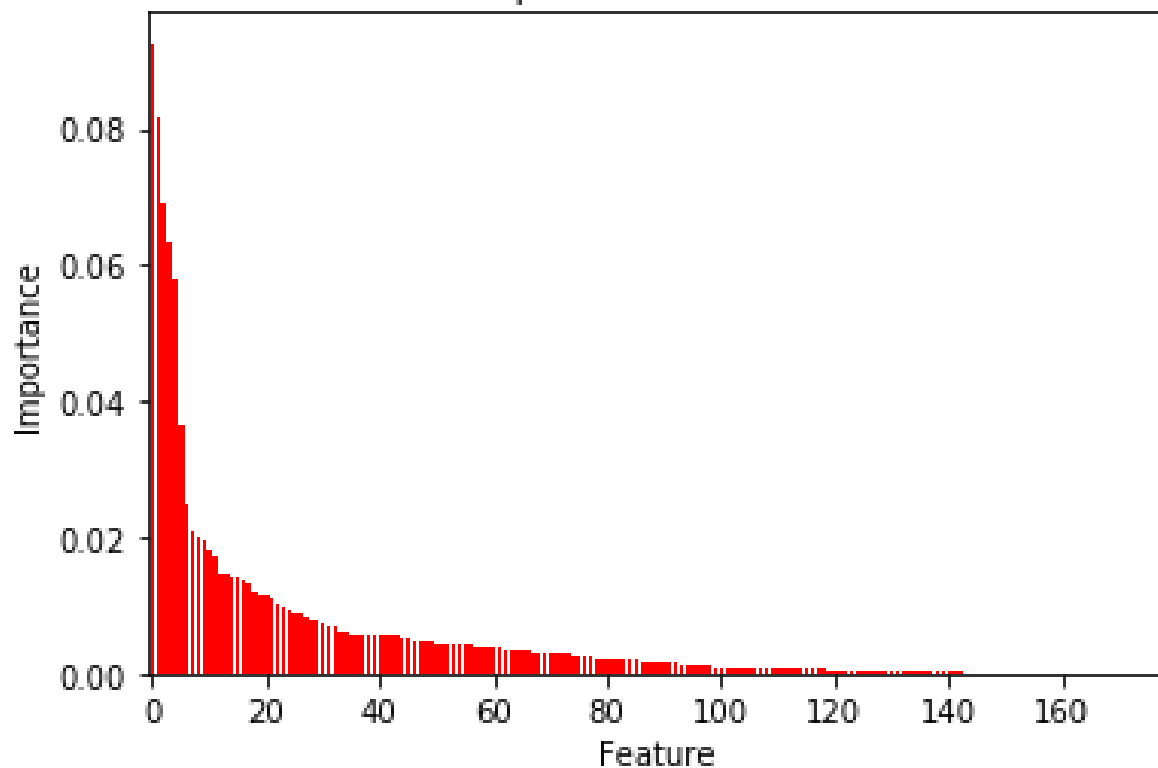
9

Model	R^2 (mean)	R^2 (variance)
Random Forest	0.41	0.11
AdaBoost	0.44	0.11

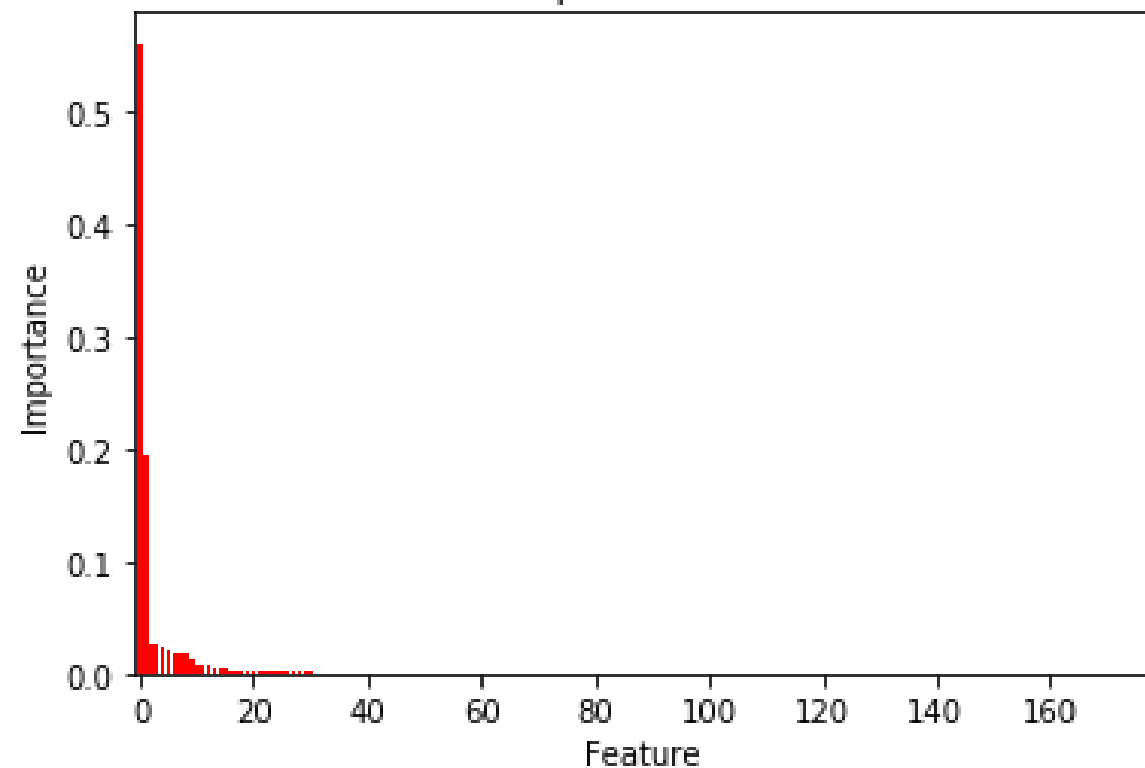
Feature Importance

10

Feature Importance (Random Forest)



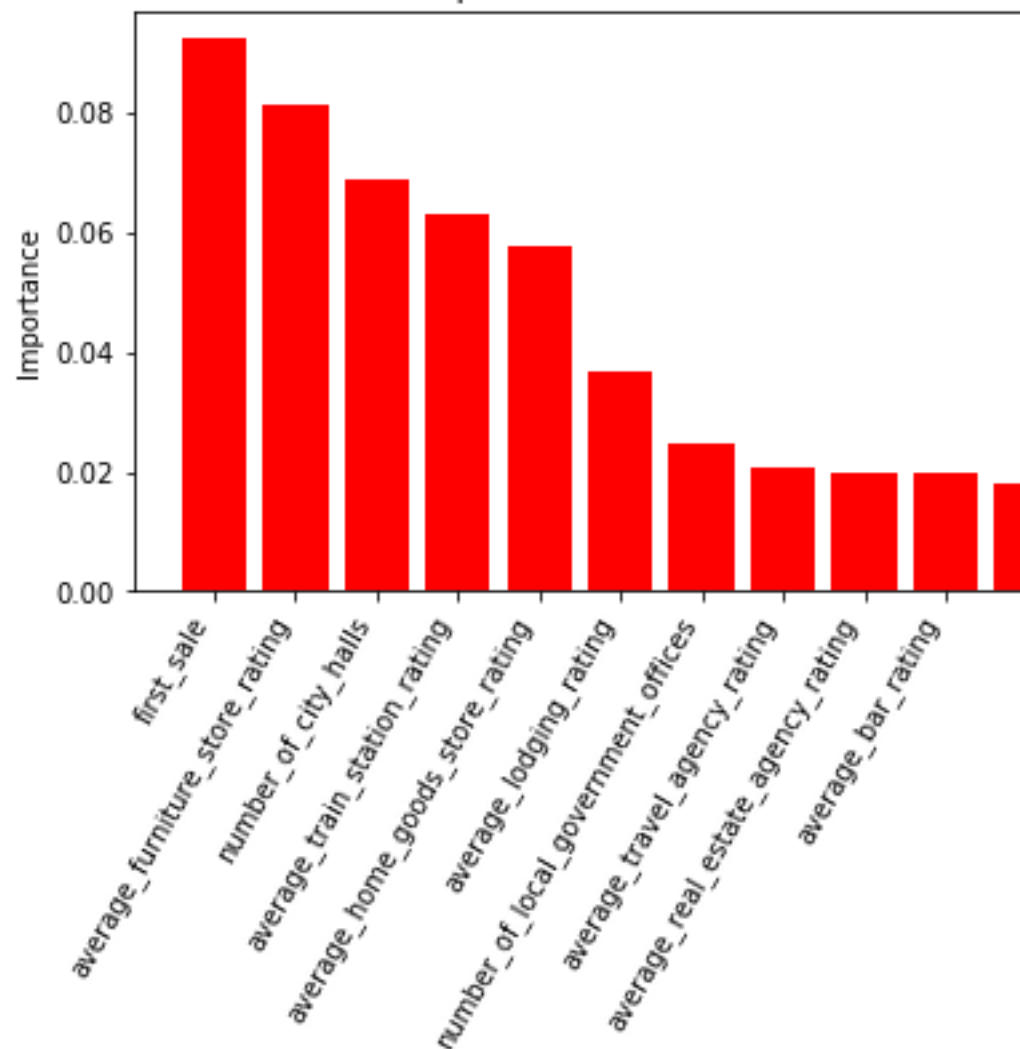
Feature Importance (AdaBoost)



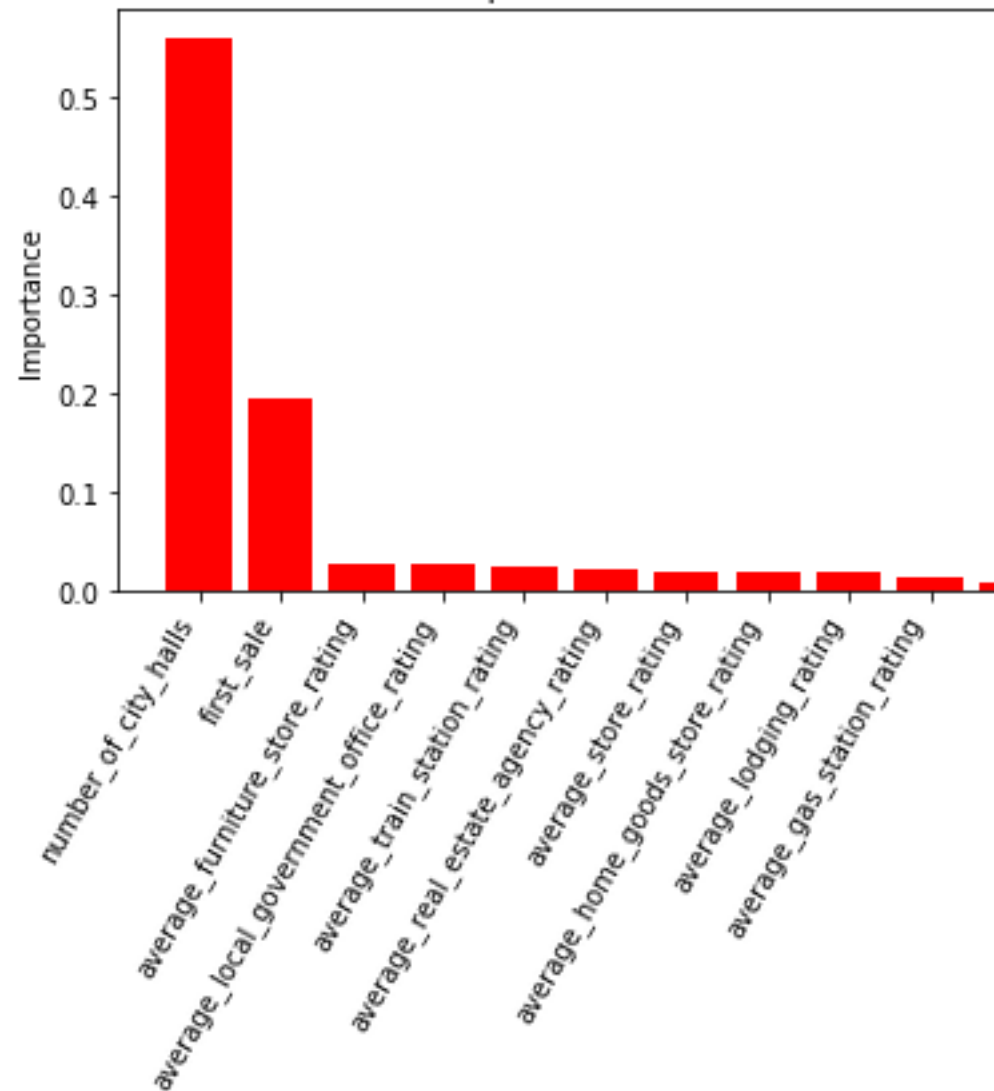
Feature Importance

11

Feature Importance (Random Forest)

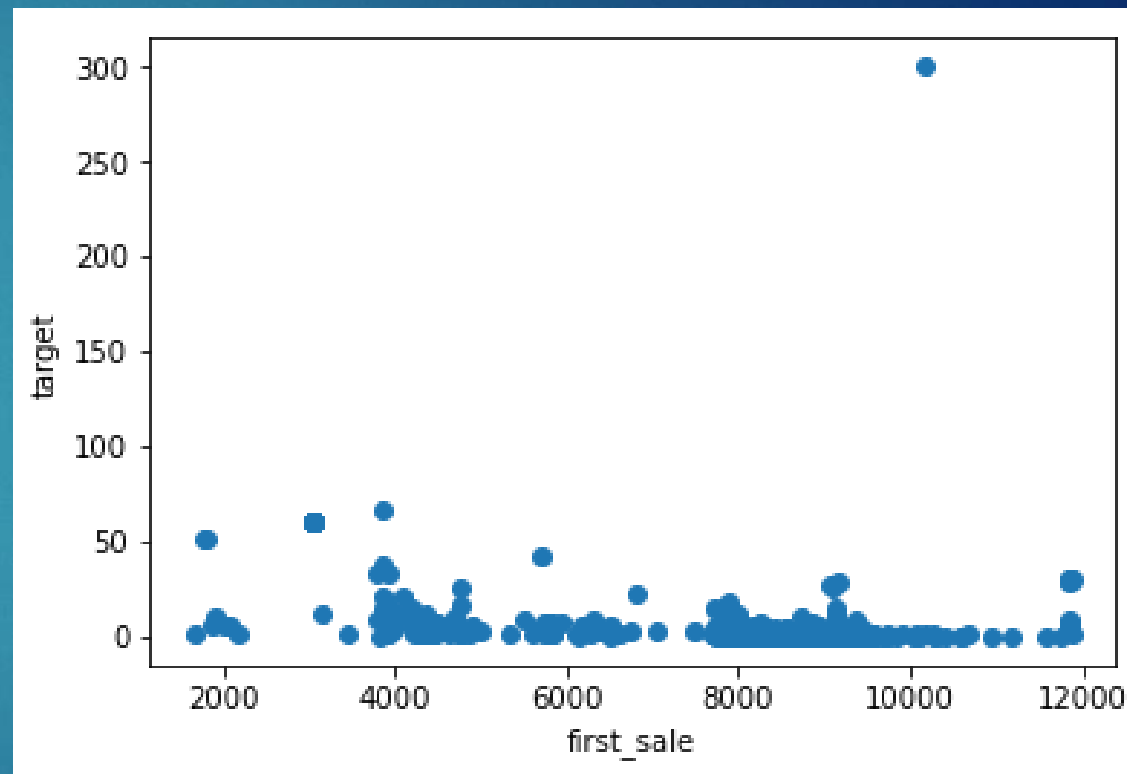
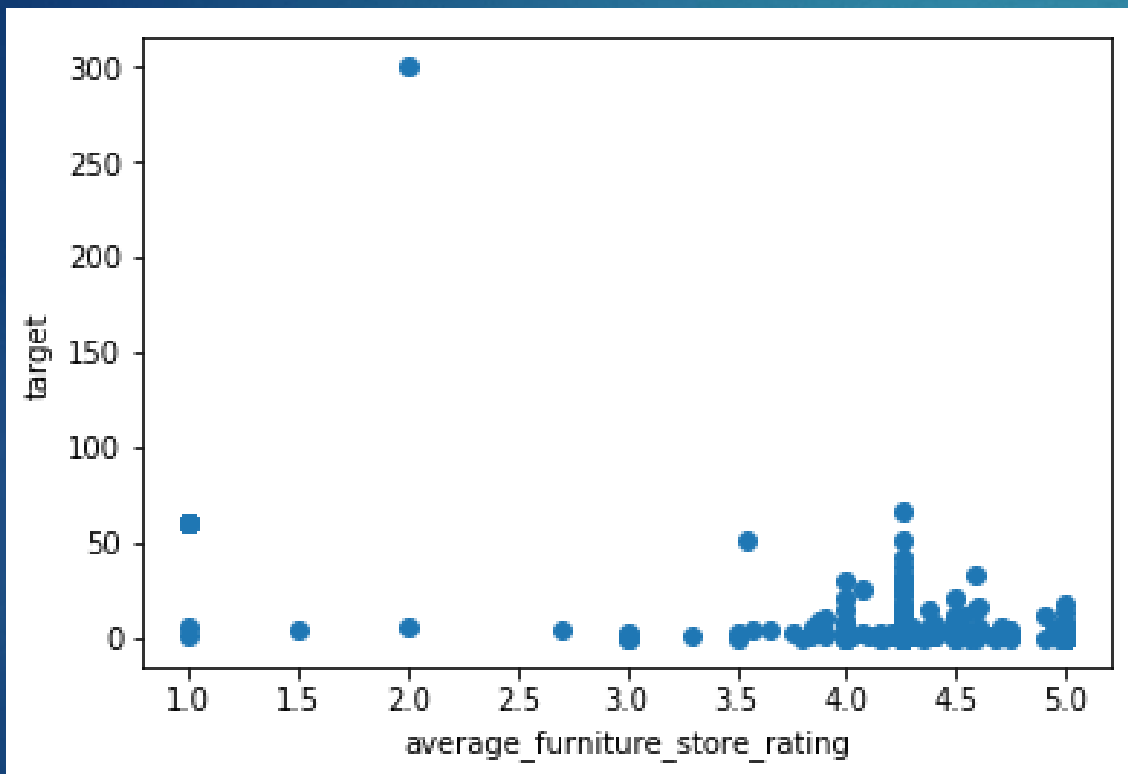


Feature Importance (AdaBoost)



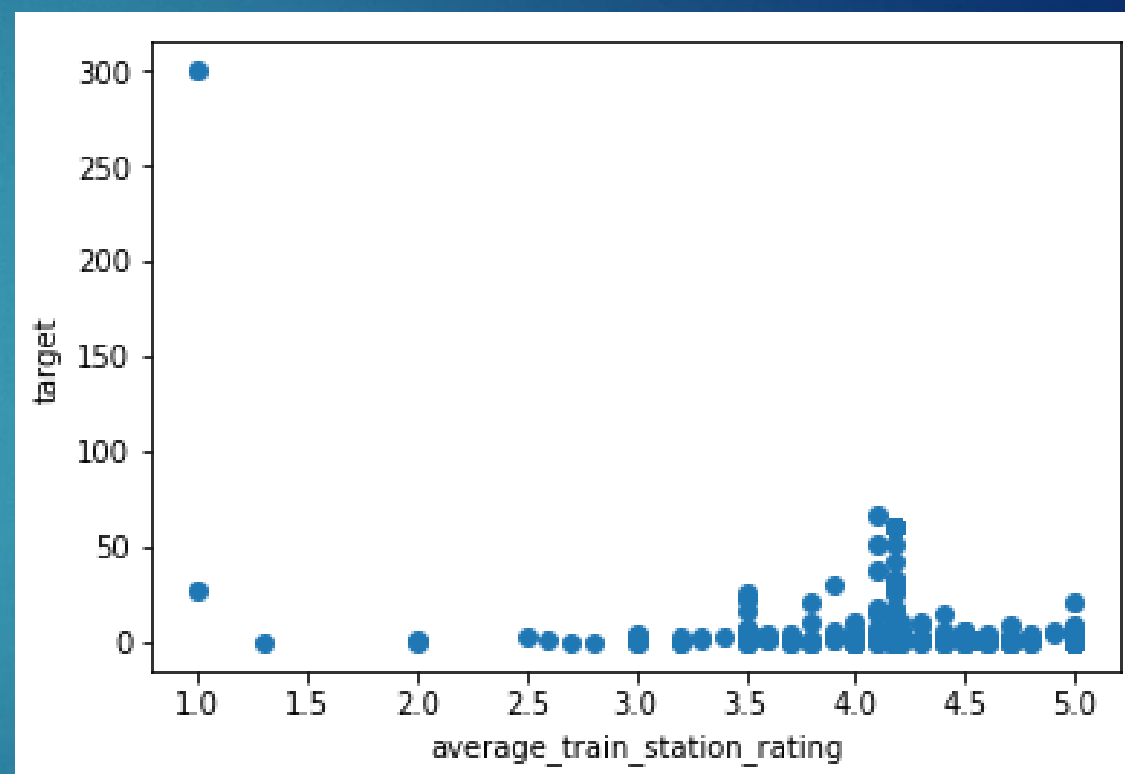
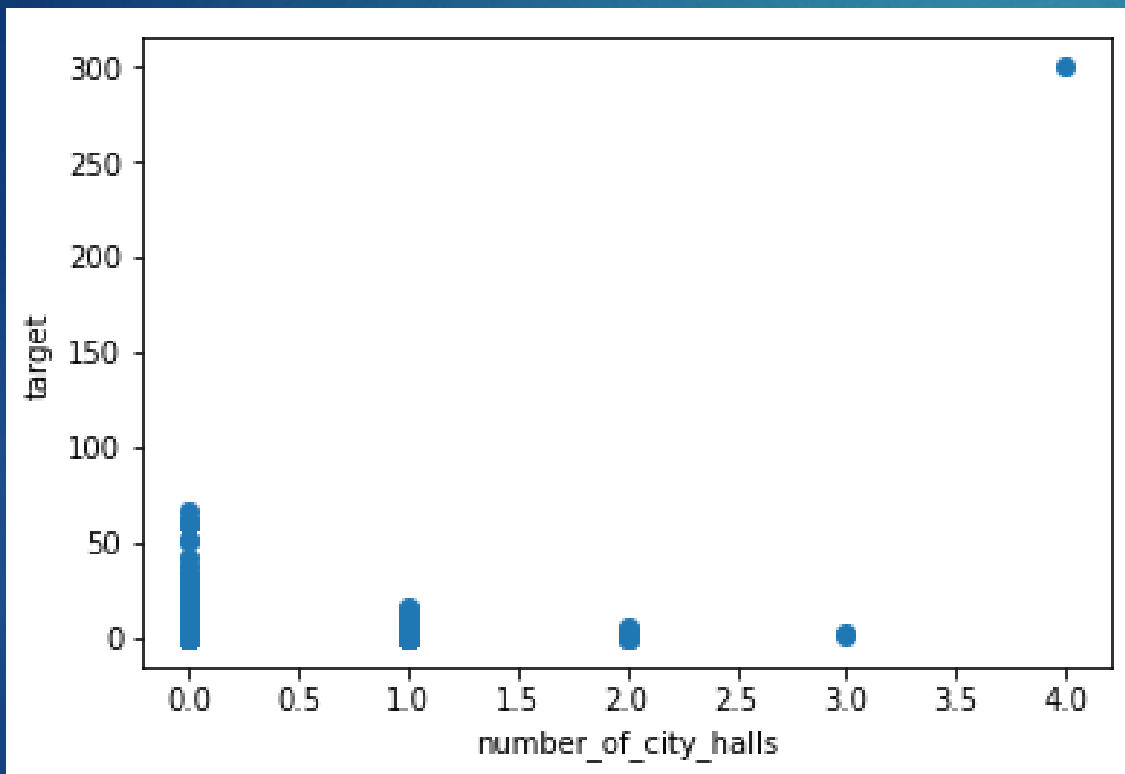
Important Features

12



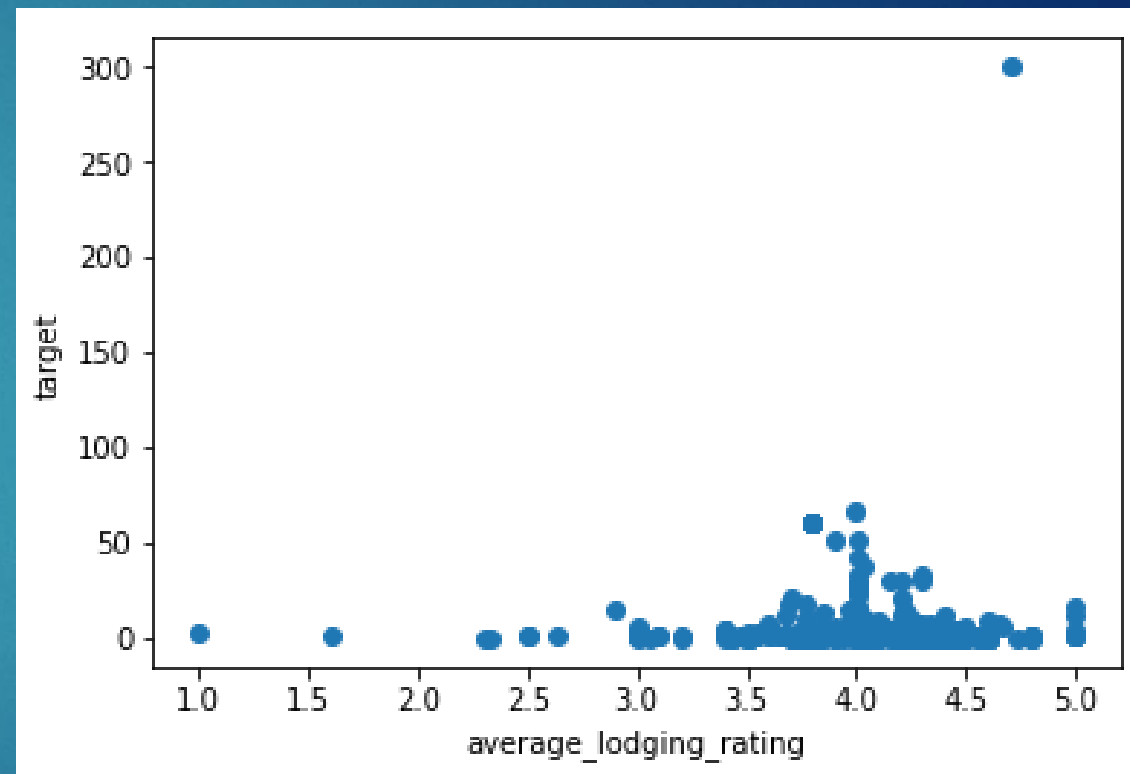
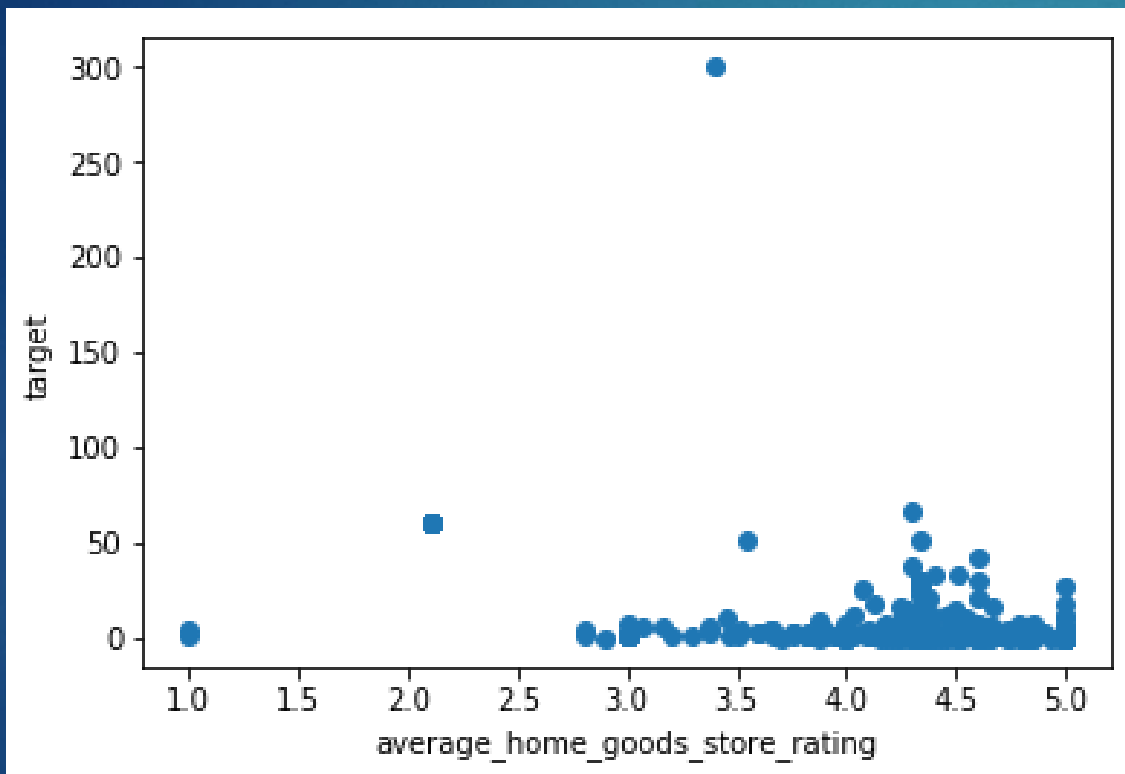
Important Features

13



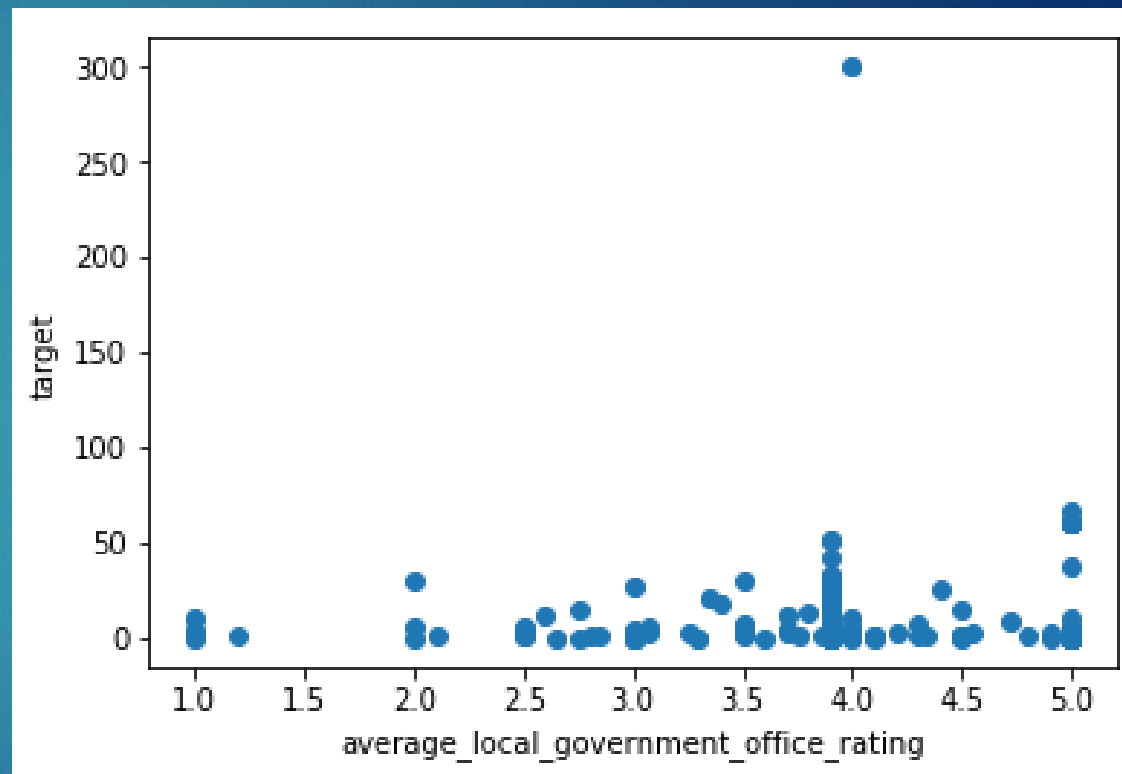
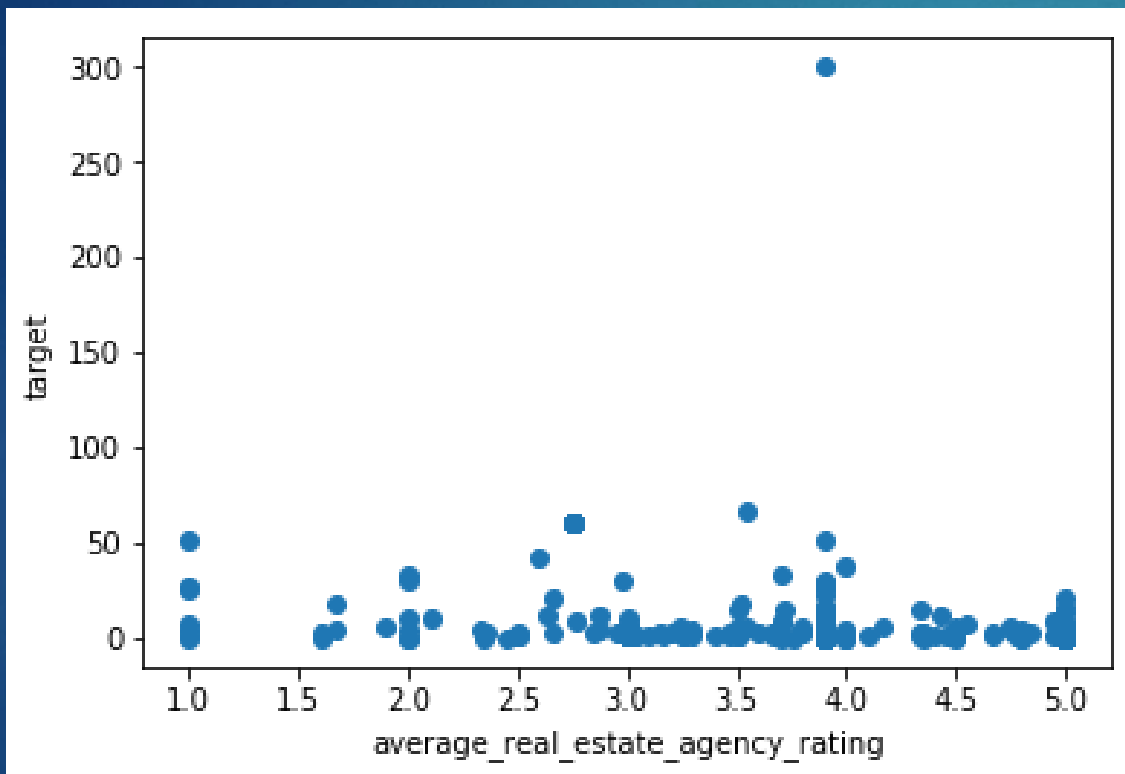
Important Features

14



Important Features

15



Conclusions

16

- ▶ Models reduce error by more than 40% (compared to baseline)
 - ▶ Surroundings influence sales only partially
- ▶ Different models find similar feature importances
 - ▶ Real estate agencies, governments, furniture stores, ...
 - ▶ Features related to houses / living / furniture ...